

**WHAT IS CLAIMED IS:**

1           1.       A method of interfacing with a machine, comprising:  
2           acquiring sets of contemporaneous images of an interactive space from  
3 multiple respective fields of view;  
4           detecting an input target in the acquired images;  
5           computing coordinates of the input target detected in the acquired images;  
6           constructing a spatiotemporal input data structure linking input target  
7 coordinates computed from contemporaneous images to respective reference  
8 times;  
9           processing the spatiotemporal input data structure to identify an input  
10 instruction; and  
11          executing the identified input instruction on the machine.

1           2.       The method of claim 1, wherein images of the interactive space are  
2 acquired from at least one stereoscopic pair of fields of view directed along  
3 substantially parallel axes intersecting the interactive space.

1           3.       The method of claim 1, wherein images of the interactive space are  
2 acquired from at least three different fields of view.

1           4.       The method of claim 1, wherein detecting the input target in the  
2 acquired images comprises comparing values of pixels in the acquired images to  
3 at least one threshold pixel value.

1           5.       The method of claim 4, wherein computing coordinates of the input  
2 target comprises computing coordinates of centroids of respective groups of pixels  
3 in the acquired images with values greater than the at least one threshold pixel  
4 value.

1           6.       The method of claim 4, wherein detecting the input target in the  
2 acquired images comprises segmenting foreground pixels from background pixels  
3 in the acquired images.

1           7.     The method of claim 1, wherein computing coordinates of the  
2 detected input target comprises computing two-dimensional coordinates of the  
3 input target detected in the acquired images.

1           8.     The method of claim 7, further comprising computing calibration  
2 parameters for the multiple fields of view.

1           9.     The method of claim 8, wherein computing coordinates of the  
2 detected input target comprises computing three-dimensional coordinates of the  
3 input target in the interactive space based on the computed two-dimensional  
4 coordinates and the computed calibration parameters.

1           10.    The method of claim 9, wherein the spatiotemporal input data  
2 structure links two-dimensional coordinates and three-dimensional coordinates  
3 computed from contemporaneous images to respective reference times.

1           11.    The method of claim 1, further comprising acquiring color values of  
2 the detected input target in the acquired images.

1           12.    The method of claim 11, wherein the spatiotemporal input data  
2 structure links input target coordinates and associated color values acquired from  
3 contemporaneous images to respective reference times.

1           13.    The method of claim 1, wherein the spatiotemporal input data  
2 structure is constructed in the form of a linked list of data records.

1           14.    The method of claim 1, wherein processing the spatiotemporal input  
2 data structure comprises identifying traces of the input target in the interactive  
3 space, each trace including a set of connected data items in the spatiotemporal  
4 input data structure.

1           15.    The method of claim 14, wherein identifying traces comprises  
2 detecting state change events and segmenting traces based on detected state  
3 change events.

1           16.     The method of claim 14, wherein identifying traces comprises  
2     computing coordinates of bounding regions encompassing respective traces.

1           17.     The method of claim 16, wherein the computed bounding region  
2     coordinates are two-dimensional coordinates of areas in the acquired images.

1           18.     The method of claim 16, wherein the computed bounding region  
2     coordinates are three-dimensional coordinates of volumes in the interactive space.

1           19.     The method of claim 14, wherein the spatiotemporal input data  
2     structure is processed to interpret the identified input target traces.

1           20.     The method of claim 19, further comprising comparing an identified  
2     trace to a predefined representation of an input gesture corresponding to a  
3     respective input instruction.

1           21.     The method of claim 20, wherein processing the spatiotemporal  
2     input data structure comprises translating the trace into a predefined  
3     alphanumeric character.

1           22.     The method of claim 19, further comprising comparing an identified  
2     trace to a location in the interactive space corresponding to a virtual interactive  
3     object.

1           23.     The method of claim 22, wherein the virtual interactive object  
2     corresponds to a virtual machine instruction input.

1           24.     The method of claim 23, wherein the virtual machine instruction  
2     input is predefined.

1           25.     The method of claim 23, further comprising constructing the virtual  
2     machine instruction input in response to processing of at least one identified input  
3     target trace.

1           26.     The method of claim 23, wherein the virtual machine instruction  
2     input corresponds to a respective mode of interpreting traces.

1           27.    The method of claim 1, wherein executing the identified input  
2 instruction comprises displaying an image in accordance with the identified input  
3 instruction.

1           28.    The method of claim 27, wherein the displayed image comprises a  
2 combination of image data generated based on the acquired images and machine-  
3 generated virtual image data.

1           29.    The method of claim 27, further comprising displaying a sequence  
2 of images at the display location showing a virtual object being manipulated in  
3 accordance with one or more identified input instructions.

1           30.    The method of claim 1, further comprising interpolating between  
2 fields of view to generate a synthetic view of the interactive space.

1           31.    A system for interfacing with a machine, comprising:  
2           multiple imaging devices configured to acquire sets of contemporaneous  
3 images of an interactive space from multiple respective fields of view; and  
4           a processing module configured to detect an input target in the acquired  
5 images, compute coordinates of the input target detected in the acquired images,  
6 construct a spatiotemporal input data structure linking input target coordinates  
7 computed from contemporaneous images to respective reference times, process  
8 the spatiotemporal input data structure to identify an input instruction, and  
9 executing the identified input instruction on the machine.

1           32.    A machine-readable medium storing machine-readable instructions  
2 for causing a machine to:  
3           acquire sets of contemporaneous images of an interactive space from  
4 multiple respective fields of view;  
5           detect an input target in the acquired images;  
6           compute coordinates of the input target detected in the acquired images;  
7           construct a spatiotemporal input data structure linking input target  
8 coordinates computed from contemporaneous images to respective reference  
9 times;

10           process the spatiotemporal input data structure to identify an input  
11   instruction; and  
12           execute the identified input instruction on the machine.

1           33.    A method of interfacing with a machine, comprising:  
2           displaying an image at a display location disposed between a viewing  
3   space and an interactive space, wherein the displayed image is viewable from a  
4   perspective in the viewing space;  
5           acquiring images of the interactive space from at least one field of view;  
6           detecting an input target in the acquired images;  
7           computing coordinates of the input target detected in the acquired images;  
8           identifying an input instruction based on the computed input coordinates;  
9   and  
10          executing the identified input instruction on the machine.

1           34.    The method of claim 33, wherein the display location corresponds  
2   to a display area of a portable electronic device.

1           35.    The method of claim 33, wherein the display location corresponds  
2   to a display area embedded in a desktop surface.

1           36.    The method of claim 33, wherein displaying the image comprises  
2   projecting the image onto a surface.

1           37.    The method of claim 33, wherein acquiring images comprises  
2   acquiring images of the interactive space from at least one field of view disposed  
3   between the display location and the interactive space.

1           38.    The method of claim 37, wherein acquiring images comprises  
2   acquiring images of the interactive space from a field of view directed toward the  
3   interactive space along an optical axis intersecting a central area of the display  
4   location.

1           39.    The method of claim 33, wherein acquiring images comprises  
2   acquiring images of the interactive space from multiple fields of view.

1           40.    The method of claim 39, further comprising interpolating between  
2 fields of view to display an image at the display location corresponding to a  
3 synthetic view of the interactive space.

1           41.    The method of claim 39, further comprising computing calibration  
2 parameters for the multiple fields of view.

1           42.    The method of claim 41, wherein computing coordinates of the  
2 detected input target comprises computing three-dimensional coordinates of the  
3 input target in the interactive space based on the computed calibration  
4 parameters.

1           43.    The method of claim 33, wherein detecting the input target in the  
2 acquired images comprises comparing values of pixels in the acquired images to  
3 at least one threshold pixel value.

1           44.    The method of claim 43, wherein computing coordinates of the  
2 input target comprises computing coordinates of centroids of respective groups of  
3 pixels in the acquired images with values greater than the threshold.

1           45.    The method of claim 43, wherein detecting the input target in the  
2 acquired images comprises segmenting foreground pixels from background pixels  
3 in the acquired images.

1           46.    The method of claim 33, wherein computing coordinates of the  
2 detected input target comprises computing two-dimensional coordinates of the  
3 input target detected in the acquired images.

1           47.    The method of claim 33, wherein identifying an input instruction  
2 comprises identifying traces of the input target in the interactive space.

1           48.    The method of claim 47, wherein identifying traces comprises  
2 detecting state change events and segmenting traces based on detected state  
3 change events.

1           49.    The method of claim 47, wherein identifying traces comprises  
2   computing coordinates of bounding regions encompassing respective traces.

1           50.    The method of claim 49, wherein the computed bounding region  
2   coordinates are two-dimensional coordinates of areas in the acquired images.

1           51.    The method of claim 49, wherein the computed bounding region  
2   coordinates are three-dimensional coordinates of volumes in the interactive space.

1           52.    The method of claim 47, wherein identifying the input instruction  
2   comprises interpreting the identified input target traces.

1           53.    The method of claim 52, further comprising comparing an identified  
2   trace to a predefined representation of an input gesture corresponding to a  
3   respective input instruction.

1           54.    The method of claim 53, wherein processing the spatiotemporal  
2   input data structure comprises translating the trace into a predefined  
3   alphanumeric character.

1           55.    The method of claim 52, further comprising comparing an identified  
2   trace to a location in the interactive space corresponding to a virtual interactive  
3   object.

1           56.    The method of claim 55, wherein the virtual interactive object  
2   corresponds to a virtual machine instruction input.

1           57.    The method of claim 56, wherein the virtual machine instruction  
2   input is predefined.

1           58.    The method of claim 56, further comprising constructing the virtual  
2   machine instruction input in response to processing of at least one identified input  
3   target trace.

1           59.    The method of claim 56, wherein the virtual machine instruction  
2   input corresponds to a respective mode of interpreting traces.

1           60.    The method of claim 33, wherein executing the identified input  
2 instruction comprises displaying an image at the display location in accordance  
3 with the identified input instruction.

1           61.    The method of claim 60, wherein the displayed image comprises a  
2 combination of image data generated based on the acquired images and machine-  
3 generated virtual image data.

1           62.    The method of claim 60, further comprising displaying a sequence  
2 of images at the display location showing a virtual object being manipulated in  
3 accordance with one or more identified input instructions.

1           63.    A system of interfacing with a machine, comprising:  
2           a display configured to present an image at a display location disposed  
3 between a viewing space and an interactive space, wherein the displayed image is  
4 viewable from a perspective in the viewing space;  
5           at least one imaging device configured to acquire images of the interactive  
6 space from at least one respective field of view; and  
7           a processing system configured to detect an input target in the acquired  
8 images, compute coordinates of the input target detected in the acquired images,  
9 identify an input instruction based on the computed input coordinates, and  
10 execute the identified input instruction on the machine.

1           64.    A machine-readable medium storing machine-readable instructions  
2 for causing a machine to:  
3           display an image at a display location disposed between a viewing space  
4 and an interactive space, wherein the displayed image is viewable from a  
5 perspective in the viewing space;  
6           acquire images of the interactive space from at least one field of view;  
7           detect an input target in the acquired images;  
8           compute coordinates of the input target detected in the acquired images;  
9           identify an input instruction based on the computed input coordinates; and  
10          execute the identified input instruction on the machine.